

PATENT APPLICATION

ASSAY CUSTOMIZATION

Inventor(s): Jonathan C. Heller, a citizen of The United States of America, residing in San Francisco, California;
Carol A. Dahl, a citizen of The United States of America, residing in Mercer Island, Washington;
John T. Stults, a citizen of The United States of America, residing in Redwood City, California;
Peter Foley, a citizen of The United States of America, residing in Los Altos Hills, California;
Stoughton L. Ellsworth, Jr., a citizen of The United States of America, residing in Palo Alto, California;
Frank Andel, III, a citizen of The United States of America, residing in Berkeley, California; and
Al Greenquist, a citizen of The United States of America, residing in San Jose, California.

Assignee: **Biospect, Inc.**
201 Gateway Boulevard
South San Francisco, CA 94080



Wilson Sonsini Goodrich & Rosati
PROFESSIONAL CORPORATION

650 Page Mill Road
Palo Alto, CA 94304
(650) 493-9300
(650) 493-6811

Express Mail Label No.: EV 333487694 US

ASSAY CUSTOMIZATION

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application is a Continuation in Part of Application Serial No. 10/645,863 filed August 20, 2003, which is claiming priority to 60/473,272 filed May 22, 2003, both incorporated herein by reference for all purposes.

BACKGROUND OF THE INVENTION

[0002] The present inventions provide a system for customization of assays such as assays based on the use of mass spectrometry.

[0003] A common aspect of all life on earth is the use of polypeptides as functional building blocks and the encryption of the instructions for the building blocks in the blueprint of nucleic acids (DNA, RNA). What distinguishes between living entities lies in the instructions encoded in the nucleic acids of the genome and the way the genome manifests itself in response to the environment as proteins. The complement of proteins, protein fragments, and peptides present at any specific moment in time defines who and what we are at that moment, as well as our state of health or disease.

[0004] One of the greatest challenges facing biomedical research and medicine is the limited ability to distinguish between specific biological states. This is reflected in the limited ability to detect the earliest stages of disease, anticipate the path any apparent disease will take in one patient versus another, predict the likelihood of response for any individual to a particular treatment, and preempt the possible adverse affects of treatments on a particular individual.

[0005] New technologies and strategies are needed to inform medical care and improve the repertoire of medical tools, as well as methods to utilize such technologies and strategies.

BRIEF SUMMARY OF THE INVENTION

[0006] According to one embodiment of the invention a method of performing analysis in a mass spectrometry system is provided. The method includes the steps of inputting a plurality of case samples and control samples; identifying a pattern of polypeptides associated with the cases and the controls; among the proteins identified in the case samples and the control

samples, selecting at least selected protein signals that are present in both the case samples and the control samples; and performing an assay on a selected sample by removing at least some of the protein represented by said signals; and among remaining proteins, associating the selected sample as associated with the cases or the controls.

BRIEF DESCRIPTION OF THE FIGURES

[0007] Figure 1 a diagram illustrating preferred aspects systems used herein.

[0008] Figure 2 illustrates a timing diagram showing operation of a parallel system.

DETAILED DESCRIPTION OF THE INVENTION

[0009] The systems herein are used to differentiate biological states with reliability, reproducibility, and sensitivity. In one embodiment, the system relies on an integrated, reproducible, sample preparation, separation, and electrospray ionization system in a microfluidics format, with high sensitivity mass spectrometry and informatics. This system will serve as the foundation for the discovery of patterns of polypeptides or other biological markers that reflect and differentiate biological states specific for various states of health and disease. For purposes herein, polypeptides includes, e.g. proteins, peptides, and/or protein fragments.

[0010] These patterns of polypeptides that reflect and differentiate biological states will then be utilized in clinically useful formats and in research contexts. Clinical applications will include detection of disease; distinguishing disease states to inform prognosis, selection of therapy, and the prediction of therapeutic response; disease staging; identification of disease processes; prediction of efficacy; prediction of adverse response; monitoring of therapy associated efficacy and toxicity; and detection of recurrence.

[0011] The system used herein may be utilized in both the applications of studying protein patterns that distinguish case and control samples, and/or in using patterns to diagnose individuals. Case samples will be those wherein a patient exhibits a particular disease state or other phenotype. For example, the case samples may be those where a patient exhibits a response to a drug. Conversely, control samples will be collected from patients that do not exhibit the phenotype under study, such as those that do not have the disease or response to a drug.

[0012] Preferably more than 10 case and 10 control samples are collected for identifying protein signals of interest. Preferably more than 20 case and 20 control samples, preferably more than 50 case and 50 control samples, preferably more than 100 case and 100 control samples, and most preferably more than 500 case and 500 control samples are collected.

[0013] The case and control samples are assayed to identify patterns of markers that are present in the case and control samples. In preferred embodiments the markers are polypeptides such as proteins, although they may also include small molecules, nucleic acids, polysaccharides, metabolites, lipids, or the like. Preferably, the patterns are obtained without advance selection or screening of the particular polypeptides involved. In some embodiments the patterns are obtained without identification of some or all of the markers that are shown in the pattern.

[0014] Preferably the assay identifies the presence of more than 100 polypeptides, preferably more than 200 polypeptides, more preferably more than 500 polypeptides, more preferably more than 1000 polypeptides, and more preferably more than 2000 polypeptides. While the identity of some of the polypeptides will be known from prior studies, it is not necessary to identify specifically all of the polypeptides indicated by the assay. The presence of (or absence of) a pattern of many polypeptides repeatedly found to be in the cases in a pattern distinct from the controls will be used in the study of phenotypes and/or diagnostics. In various embodiments a number of polypeptides are represented in the pattern, but the identity of some of these polypeptides is not known. For example, more than 15 polypeptides can be represented, more than 30 polypeptides can be represented, more than 50 polypeptides can be represented, more than 100 polypeptides can be represented, and more than 1000 polypeptides can be represented.

[0015] Preferably such systems allow for the capture and measure of many or all of the instances of a polypeptide in a sample that is introduced in the mass spectrometer for analysis. Using such systems it is preferable that one can observe those polypeptides with high information-content but that are only present at low concentrations, such as those “leaked” from diseased tissue. Other high information-content polypeptides may be those that are related to the disease, for instance, those that are generated in the tumor-host environment.

[0016] In some embodiments, an early assay, or discovery experiment, is followed by a later assay. The early assay will normally be used in initial identification of the polypeptides that identify or separate cases from controls. The later assay is adjusted according to parameters

that can focus diagnostics or evaluation of regions of interest, such as regions of high differentiation, i.e. those regions or markers where there are significant differences between case samples and control samples. The parameters can be determined by, for example, an early assay which may identify the regions of interest, which may be on one technology platform, and a later assay on the same or a different platform.

[0017] A bioinformatics system is utilized to identify the differences in the polypeptide patterns in the case and control samples. Patterns will be composed of the relative representation of numerous polypeptides or other biological entities, the collective profile of which will be more important than the presence or absence of any specific entities. By identifying patterns in blood or other patient samples, the methods will not only provide the window to the presence of disease and other pathology in some embodiments, but also to the body's ongoing response to the disease or pathologic condition in other embodiments. In a high throughput mode, data from a first sample are evaluated in a bio-informatics system at the same time another sample is being processed in, for example, a mass spectrometry system.

[0018] The patterns of polypeptides present in the sample may be used to identify the disease state of a patient sample in, for example, a diagnostic setting. Samples will, in preferred embodiments be serum samples, although tissue or bodily fluid samples from a variety of sources will be used in alternative embodiments. Preferably, though not necessarily, the system used in the diagnostic application is based upon the same technology platform as the platform used to identify the patterns in the first instance. For example, if the platform used to identify the patterns in the first instance is a time of flight (TOF) mass spectrometer, it is preferred that the diagnostic applications of the patterns are run on a time of flight mass spectrometer.

[0019] In preferred embodiments, the mass spectrometer utilized herein is coupled to a microfluidic separations device. The sample preparation techniques used thereon preferably concentrate the polypeptides the mass spectrometer is best able to detect and/or are which are most informative, and deplete the ones that are more difficult to detect and/or are less informative (because, for example, they appear in both case and control samples).

[0020] In most preferred embodiments the microfluidic separations device is a disposable device that is readily attached to and removed from the mass spectrometer, and sold as a disposable, thereby providing a recurring revenue stream to the involved business and a reliable product to the consumer. Preferably, a mass spectrometer is utilized that will accept a

continuous sample stream for analysis and provide high sensitivity throughout the detection process.

[0021] Sample preparation will, in some embodiments, include the removal of high abundance polypeptides, denaturation, removal of polypeptides expected to be in abundance in all samples, addition of preservatives and calibrants, and desalting. These steps will allow sensitive measurement of concentrations of information-rich polypeptides, such as those that have leaked from tissue, as compared to polypeptides that would carry little information, such as those highly abundant and native to serum. Prepared samples will then be separated using fast molecular separations methods with high peak capacities. An electrospray ionization (ESI) interface may be integrated on the microfluidics chip, which will ionize and spray the prepared and separated serum directly into a mass spectrometer and is preferably sold as part of a disposable component to assure high reliability of the system.

[0022] The microfluidics-based separations preferably provide the polypeptide mixtures at flow rates and at complexity levels that are matched to the mass spectrometer's optimal performance regions. The mass spectrometer's sensitivity is preferably optimized to detect the species most likely to differentiate biological states. Preferably, the reagents necessary for performing these steps are provided in or along with the microfluidics device, thereby allowing for additional recurring revenue to the involved business and higher performance for the user.

[0023] The sample preparation system will provide for different operations depending upon the detection device to be utilized. The sample preparation system preferably provides for protein denaturation prior to processing on the mass spectrometer. Analytes of interest herein will in some cases be in a protein bound form. Preferably the system provides for denaturation of proteins preferably prior to the removal of high abundance materials (such as albumin or other proteins from serum or plasma samples). By denaturing such proteins prior to their removal, bound analytes of interest will be released such that they can be meaningful in later analysis. Denaturation may utilize any of several techniques including the use of heat, high salt concentrations, the use of acids, base, chaotropic agents, organic solvents, detergents and/or reducing agents. Liotta, Lance, A., et al., "Written in Blood," *Nature* (10/30/2003), Volume 425, page 905. Tirumalai, Radhakrishna S., et al. "Characterization of the Low Molecular Weight Human Serum Proteome," *Molecular & Cellular Proteomics* 2.10 (8/13/2003), pages 1096-1103.

[0024] The system used for removal of high abundance polypeptides may be based on, for example, the use of high affinity reagents for removal of the polypeptides, the use of high molecular weight filters, ultracentrifugation, precipitation, and/or electrodialysis.

Polypeptides that will often be removed will include, for example, those involved in normal metabolism, and a wide variety of other indications not of relevance to a particular assay.

Such proteins may be removed through, for example, a solid phase extraction resin.

Additionally, the system may include a reversed phase chromatography device, for example, for separation of small molecules and/or to trap, desalt, and separate a protein mixture.

[0025] Figure 1 illustrates additional aspects of an exemplary system platform used herein.

The invention involves an integrated system to a) discover; and b) assay patterns of polypeptides that reflect and differentiate biological and clinical states of organisms, including patients, in biological materials including but not limited to body fluids. Biological and clinical states include but are not limited to states of development; age; health; pathology; disease detection, process, or staging; infection; toxicity; or response to chemical, environmental, or drug factors (such as drug response phenotyping, drug toxicity phenotyping, or drug effectiveness phenotyping). Biological fluids 201 include but are not limited to serum, plasma, whole blood, nipple aspirate, pancreatic fluid, trabecular fluid, lung lavage, urine, cerebrospinal fluid, saliva, sweat, pericrevicular fluid, and tears.

[0026] The system provides for the integration of fast molecular separations and electrospray ionization system 204 on a microfluidics platform 203. The system provides processed samples to a high sensitivity time of flight mass spectrometer 205. Signal processing system and pattern extraction and recognition tools 205 incorporate domain knowledge to extract information from polypeptide patterns and classify the patterns to provide a classification 209.

The signal processing system may include or be coupled to other software elements as well. For example, the signal processing system may provide for an easy to use user interface on the associated computer system and/or a patient database for integration of results into an institution's laboratory or patient information database system.

[0027] The microfluidics device(s) 203 may be formed in plastic by means of etching, machining, cutting, molding, casting or embossing. The microfluidics device(s) may be made from glass or silicon by means of etching, machining, or cutting. The device may be formed by polymerization on a form or other mold. The molecular separations unit or the integrated fast molecular separations/electrospray ionization unit may provide additional sample

preparation steps, including sample loading, sample concentration, removal of salts and other compounds that may interfere with electrospray ionization, removal of highly abundant species, concentration of the sample to a smaller volume, proteolytic or chemical cleavage of components within the biological material, enzymatic digestion, and/or aliquoting in to storage containers. The particular operations performed by the device will depend upon the detection technology that is utilized.

[0028] The device(s) for separations and electrospray may be either single use for a single sample, multi-use for a single sample at a time with serial loading, single use with parallel multiple sample processing, multi-use with parallel multiple sample processing or a combination. Separations processes may include isoelectric focusing, electrophoresis, chromatography, or electrochromatography. The separations device may include collection areas or entities for some or all of the purified or partially purified fractions.

[0029] It is to be understood that the inventions herein are illustrated primarily with regard to mass spectrometry as a detection device, but other devices may be used alone or with the mass spectrometer. For example, detection devices may include electrochemical, spectroscopic, or luminescent detectors, and may be integral with the microfluidics device.

[0030] Mass spectrometers that may be used include quadrupole, ion trap, magnetic sector, Fourier transform ion cyclotron resonance instruments, or an orthogonal time-of-flight mass spectrometer which includes an analyzer that receives an ion beam from an electrospray ionization (ESI) source.

[0031] In preferred embodiments the system also adapts the speed of the system in response to the detection of known markers that are likely to be present in all samples, and which are readily detectable. Since separations will often vary in retention or migration time, by detecting molecules that are known, likely to be in all samples, and easily detectable, and then comparing the speed at which they have passed through the system in comparison to a standard from other experiments, it becomes possible to speed the system up by speeding the separations in response to the detection of slower than expected migration time, or slowing the system down in response to faster than expected migration times. The speed may be adjusted through, for example, adjustments in system pressure, voltage, current flow, or temperature. Preferably, the system is operated faster or slower by changing the voltage. Representative peptides and proteins that could be spiked into samples and could be used for this purpose include neurotensin, lysozyme, aprotinin, insulin b-chain, and renin substrate. In

addition, the speed of operation of the device may be slowed to provide greater accuracy in the detection of molecules of particular interest in a spectrum. Conversely, the system may be operated more quickly during the times when components of low interest would be expected to be detected.

[0032] In some embodiments pressure is added to move the components through the electrophoretic device, especially to migrate components to the end of an electrophoretic separation capillary (in conjunction with the use of the electro osmotic flow). The pressure produces buffer flow that is required to maintain a stable electrospray.

[0033] Ions formed by electrospray ionization will normally be chiefly singly or multiply charge ions of molecules, with charge coming from protons or alkali metal bound to the molecules. Ion excitation may be produced by collision of ions with background gas or an introduced collision gas. Alternatively, excitation may be from collision with other ions, a surface, interaction with photons, heat, electrons, or alpha particles. Through excitation of the sample in an electrospray the information content of the process should be altered and/or enhanced. Such excitation may, for example, desolvate ions, dissociate noncovalently bound molecules from analyte ions, break up solvent clusters, fragment background ions to change their mass to charge ratio and move them to a ratio that may interfere less with the analysis, strip protons and other charge carriers such that multiply charged ions move to different regions of the spectrum, and fragment analyte ions to produce additional, more specific or sequence-related information.

[0034] In preferred embodiments the excitation system may be turned on and off to obtain a set of spectra in both states. The information content of the two spectra will, in most cases, be far greater than the information content of either single spectra. In such embodiments the system will include a switching device for activating and de-activating the excitation/ionization system. Analysis software will be configured in this case to analyze the sample separately both in the “on” state of the excitation system and in the “off” state of the excitation system. Different markers may be detected more efficiently in one or the other of these two states.

[0035] Fig. 2 illustrates the pipelined systems operations in greater detail. As shown at step 351, a first sample is acquired during this time frame and separated in the microfluidics device, and then processed in the mass spectrometer. At step 353 a second sample is processed in the microfluidics device and processed in the mass spectrometer. During at least

some of the time when second sample is being processed at step 353, the data from the mass spectrum for the first sample are processed in the data analysis system at step 357. Similarly, at step 355 a third sample is processed in the microfluidics device and the mass spectrometer, while the data from sample 2 are being analyzed in the data analysis system at step 359.

Sample Collection

[0036] Case samples are obtained from individuals with a particular phenotypic state of interest. Examples of phenotypic states include, phenotypes resulting from an altered environment, drug treatment, genetic manipulations or mutations, injury, change in diet, aging, or any other characteristic(s) of a single organism or a class or subclass of organisms. In a preferred embodiment, a phenotypic state of interest is a clinically diagnosed disease state. Such disease states include, for example, cancer, cardiovascular disease, inflammatory disease, infectious disease and pregnancy related disorders. Control samples are obtained from individuals who do not exhibit the phenotypic state of interest or disease state (e.g., an individual who is not affected by a disease or who does not experience negative side effects in response to a given drug). Alternatively, states of health can be analyzed.

[0037] Cancer phenotypes are studied in some aspects of the invention. Examples of cancer studies herein include, but are not limited to, breast cancer, skin cancer, bone cancer, prostate cancer, liver cancer, lung cancer, brain cancer, cancer of the larynx, gallbladder, pancreas, rectum, parathyroid, thyroid, adrenal, neural tissue, head and neck, colon, stomach, bronchi, kidneys, basal cell carcinoma, squamous cell carcinoma of both ulcerating and papillary type, metastatic skin carcinoma, osteo sarcoma, Ewing's sarcoma, veticulum cell sarcoma, myeloma, giant cell tumor, small-cell lung tumor, gallstones, islet cell tumor, primary brain tumor, acute and chronic lymphocytic and granulocytic tumors, hairy-cell tumor, adenoma, hyperplasia, medullary carcinoma, pheochromocytoma, mucosal neuronms, intestinal ganglioneuromas, hyperplastic corneal nerve tumor, marfanoid habitus tumor, Wilm's tumor, seminoma, ovarian tumor, leiomyomater tumor, cervical dysplasia and in situ carcinoma, neuroblastoma, retinoblastoma, soft tissue sarcoma, malignant carcinoid, topical skin lesion, mycosis fungoide, rhabdomyosarcoma, Kaposi's sarcoma, osteogenic and other sarcoma, malignant hypercalcemia, renal cell tumor, polycythermia vera, adenocarcinoma, glioblastoma multiforma, leukemias, lymphomas, malignant melanomas, epidermoid carcinomas, and other carcinomas and sarcomas.

[0038] Cardiovascular disease may be studied in other applications of the invention. Examples of cardiovascular disease include, but are not limited to, congestive heart failure, high blood pressure, arrhythmias, atherosclerosis, cholesterol, Wolff-Parkinson-White Syndrome, long QT syndrome, angina pectoris, tachycardia, bradycardia, atrial fibrillation, ventricular fibrillation, congestive heart failure, myocardial ischemia, myocardial infarction, cardiac tamponade, myocarditis, pericarditis, arrhythmogenic right ventricular dysplasia, hypertrophic cardiomyopathy, Williams syndrome, heart valve diseases, endocarditis, bacterial, pulmonary atresia, aortic valve stenosis, Raynaud's disease, Raynaud's disease, cholesterol embolism, Wallenberg syndrome, Hippel-Lindau disease, and telangiectasis.

[0039] Inflammatory disease may be studied in other applications of the system. Examples of inflammatory disease include, but are not limited to, rheumatoid, arthritis, non-specific arthritis, inflammatory disease of the larynx, inflammatory bowel disorder, pelvic inflammatory disease, inflammatory disease of the central nervous system, temporal arteritis, polymyalgia rheumatica, ankylosing spondylitis, polyarteritis nodosa, Reiter's syndrome, scleroderma, systemis lupus and erythematosus.

[0040] Infectious disease may be studied in still further aspects of the system. Examples of infectious disease include, but are not limited to, AIDS, hepatitis C, SARS, tuberculosis, sexually transmitted diseases, leprosay, lyme disease, malaria, measles, meningitis, mononucleosis, whooping cough, yellow fever, tetanus, arboviral encephalitis, and other bacterial, viral, fungal or helminthic diseases.

[0041] Pregnancy related disorders include pre-eclampsia, eclampsia pre-term birth, growth restriction in utero, rhesus incompatibility, retained placenta, septicemia, separation of the placenta, ectopic pregnancy, hypermosis gravidarum, placenta previa, erythroblastosis fetalis, pruritic urticarial papula and plaques.

[0042] In some instances, samples may be collected from individuals over a longitudinal period of time (e.g., once a day, once a week, once a month, biannually or annually). The longitudinal period may, for example, also be before, during, and after a stress test or a drug treatment. Obtaining numerous samples from an individual over a period of time can be used to verify results from earlier detections and/or to identify an alteration in polypeptide pattern as a result of, for example, aging, drug treatment, pathology, etc. Samples can be obtained from humans or non-humans. In a preferred embodiment, samples are obtained from humans.

[0043] Sample preparation and separation can involve any of the following procedures, depending on the type of sample collected and/or types of protein searched: removal of high abundance polypeptides (e.g., albumin, and transferrin); addition of preservatives and calibrants, denaturation, desalting of samples; concentration of sample polypeptides; protein digestions; and fraction collection. Preferably, sample preparation techniques concentrate information-rich polypeptides (e.g., polypeptides that have “leaked” from diseased cells or are produced by the host response to the tumor) and deplete polypeptides that would carry little or no information such as those that are highly abundant.

[0044] Sample preparation can take place in a manifold or preparation/separation device. In preferred embodiment, such preparation/separation device is a microfluidics device. Optimally, the preparation/separation device interfaces directly or indirectly with a detection device. In another embodiment, such preparation/separation device is a fluidics device.

[0045] Approximately 100 μ L of a sample or less is analyzed per assay in some particular embodiments of the invention. Removal of undesired polypeptides (e.g., high abundance, uninformative, or undetectable polypeptides) can be achieved using high affinity reagents, high molecular weight filters, ultracentrifugation and/or electrodialysis. High affinity reagents include antibodies or aptamers that selectively bind to high abundance polypeptides or reagents that have a specific pH, ionic value, or detergent strength. High molecular weight filters include membranes that separate molecules on the basis of size and molecular weight. Such filters may further employ reverse osmosis, dialysis, nanofiltration, ultrafiltration and microfiltration.

[0046] Ultracentrifugation is another method for removing undesired polypeptides. Ultracentrifugation is the centrifugation of a sample at about 60,000 rpm while monitoring with an optical system the sedimentation (or lack thereof) of particles. Finally, electrodialysis is an electromembrane process in which ions are transported through ion permeable membranes from one solution to another under the influence of a potential gradient. Since the membranes used in electrodialysis have the ability to selectively transport ions having positive or negative charge and reject ions of the opposite charge, electrodialysis is useful for concentration, removal, or separation of electrolytes.

[0047] In a preferred embodiment, the manifold or microfluidics device performs electrodialysis to remove high molecular weight polypeptides or undesired polypeptides. Electrodialysis is first used to allow only molecules under approximately 30 kD (not a sharp

cutoff) to pass through into a second chamber. A second membrane with a very small molecular weight (roughly 500 D) will allow smaller molecules such as salts to egress the second chamber.

[0048] After samples are prepared, polypeptides of interest may be separated. Separation can take place in the same location as the preparation or in another location. In a preferred embodiment, separation occurs in the same microfluidics device where preparation occurs, but in a different location on the device. Samples can be removed from an initial manifold location to a microfluidics device using various means, including an electric field. In one embodiment, the samples are concentrated during their migration to the microfluidics device using reverse phase beads and an organic solvent elution such as 50% methanol. This elutes the molecules into a channel or a well on a separation device of a microfluidics device. In another embodiment, samples are concentrated by isotachopheresis, in which ions are concentrated at a boundary between a leading and a trailing electrolyte of lower and higher electrophoretic mobilities, respectively.

[0049] Separation can involve any procedure known in the art, such as capillary electrophoresis (e.g., in capillary or on-chip), or chromatography (e.g., in capillary, column or on a chip).

[0050] Electrophoresis is the separation of ionic molecules such as polypeptides by differential migration patterns through a gel based on the size and ionic charge of the molecules in an electric field. Electrophoresis can be conducted in a gel, capillary or on a chip. Examples of gels used for electrophoresis include starch, acrylamide, agarose or combinations thereof. In a preferred embodiment, polyacrylamide gels are used. A gel can be modified by its cross-linking, addition of detergents, immobilization of enzymes or antibodies (affinity electrophoresis) or substrates (zymography) and pH gradient. Examples of capillaries used for electrophoresis include capillaries that interface with an electrospray.

[0051] Capillary electrophoresis (CE) is preferred for separating complex hydrophilic molecules and highly charged solutes. Advantages of CE include its use of small samples (sizes ranging from 0.001 to 10 μ L), fast separation, easily reproducible, and the ability to be coupled to a mass spectrometer. CE technology, in general, relates to separation techniques that use narrow bore fused-silica capillaries to separate a complex array of large and small molecules. High voltages are used to separate molecules based on differences in charge, size and hydrophobicity. Depending on the types of capillary and buffers used, CE can be further

segmented into separation techniques such as capillary zone electrophoresis (CZE), capillary isoelectric focusing (CIEF) and capillary electrochromatography (CEC).

[0052] Capillary zone electrophoresis (CZE), also known as free-solution CE (FSCE), is the simplest form of CE. The separation mechanism of CZE is based on differences in the charge-to-mass ratio of the analytes. Fundamental to CZE are homogeneity of the buffer solution and constant field strength throughout the length of the capillary. The separation relies principally on the pH-controlled dissociation of acidic groups on the solute or the protonation of basic functions on the solute.

[0053] Capillary isoelectric focusing (CIEF) allows amphoteric molecules, such as polypeptides, to be separated by electrophoresis in a pH gradient generated between the cathode and anode. A solute will migrate to a point where its net charge is zero. At this isoelectric point (the solute's pI), migration stops and the sample is focused into a tight zone. In CIEF, once a solute has focused at its pI, the zone is mobilized past the detector by either pressure or chemical means.

[0054] CEC is a hybrid technique between traditional liquid chromatography (HPLC) and CE. In essence, CE capillaries are packed with HPLC packing and a voltage is applied across the packed capillary, which generates an electro-osmotic flow (EOF). The EOF transports solutes along the capillary towards a detector. Both differential partitioning and electrophoretic migration of the solutes occurs during their transportation towards the detector, which leads to CEC separations. It is therefore possible to obtain unique separation selectivities using CEC compared to both HPLC and CE. The beneficial flow profile of EOF reduces flow related band broadening and separation efficiencies of several hundred thousand plates per meter are often obtained in CEC. CEC also makes it is possible to use small-diameter packings and achieve very high efficiencies.

[0055] Chromatography is another method for separating a subset of polypeptides. Chromatography is based on the differential absorption and elution of certain polypeptides. Liquid chromatography (LC), for example, involves the use of fluid carrier over a stationary phase. Conventional LC columns have an inner diameter of roughly 4.6 mm and a flow rate of roughly 1 ml/min. Micro-LC has an inner diameter of roughly 1.0 mm and a flow rate of roughly 40 ul/min. Capillary LC utilizes a capillary with an inner diameter of roughly 300 um and a flow rate of approximately 5 ul/min. Nano-LC is available with an inner diameter of 50 um -1 mm and flow rates of 200 nl/min. Nano-LC can vary in length (e.g., 5, 15, or 25

cm) and have typical packing of C18, 5 μ m particle size. Nano-LC stationary phase may also be a monolithic material, such as a polymeric monolith or a sol-gel monolith. In a preferred embodiment, nano-LC is used. Nano-LC provides increased sensitivity due to lower dilution of chromatographic sample. The sensitivity of nano-LC as compared to HPLC can be as much as 3700 fold.

[0056] In preferred embodiments, the samples are separated on using capillary electrophoresis separation, more preferably CEC with sol-gels, or more preferably CZE. This will separate the molecules based on their electrophoretic mobility at a given pH (or hydrophobicity in the case of CEC).

[0057] In other preferred embodiments, the steps of sample preparation and separation are combined using microfluidics technology. A microfluidic device is a device that can transport liquids including various reagents such as analytes and elutions between different locations using microchannel structures. Microfluidic devices provide advantageous miniaturization, automation and integration of a large number of different types of analytical operations. For example, continuous flow microfluidic devices have been developed that perform serial assays on extremely large numbers of different chemical compounds. Microfluidic devices may also provide the feature of disposability, to prevent sample carry-over. By microfluidics device it is intended to mean herein devices with channels smaller than 1000 μ m, preferably less than 500 μ m, and more preferably less than 100 μ m. Preferably such devices use sample volumes of less than 1000 μ l, preferably less than 500 μ l, and most preferably less than 100 μ l.

[0058] In a preferred embodiment, microfluidic devices are composed of plastic and formed by means of etching, machining, cutting, molding, casting or embossing. The microfluidics devices may alternatively be made from glass or silicon by means of etching, machining, cutting, or embossing. The microfluidic devices may be either single use for a single sample; multi-use for a single sample at a time with serial loading; single use with parallel multiple sample processing; multi-use with parallel multiple sample processing; or a combination. Furthermore, more than one microfluidics device may be integrated into the system and interface with a single detection device.

[0059] Once prepared and separated, the polypeptides are automatically delivered to a detection device, which detects the polypeptides in a sample. In a preferred embodiment, polypeptides in solution are delivered to a detection device by electrospray ionization (ESI).

ESI operates by infusing a liquid containing the sample of interest through a channel or needle, which is kept at a potential (typically 3.5 kV). The voltage on the needle causes the spray to be charged as it is nebulized. The resultant droplets evaporate at atmospheric pressure or in a region maintained at a vacuum as low as several torr, until the solvent is essentially completely stripped off, leaving a charged ion. The charged ions are then detected by a detection device such as a mass spectrometer. In a more preferred embodiment, nanoelectrospray ionization is used. Nanospray ionization is a miniaturized version of ESI and provides low detection limits using extremely limited volumes of sample fluid.

[0060] In preferred embodiments, separated polypeptides are directed down a channel that leads to an electrospray ionization emitter, which is built into a microfluidic device (an integrated ESI microfluidic device). Preferably, such integrated ESI microfluidic device provides the detection device with samples at flow rates and complexity levels that are optimal for detection. Such flow rates are, preferably, approximately 50-200 $\mu\text{L}/\text{min}$. Furthermore, a microfluidic device is preferably aligned with a detection device for optimal sample capture. For example, using dynamic feedback circuitry, a microfluidic device may allow for control positioning of an electrospray voltage and for the entire spray to be captured by the detection device orifice. The microfluidic device can be sold separately or in combination with other reagents, software tools and/or devices.

[0061] Calibrants can also be sprayed into detection device. Calibrants are used to set instrument parameters and for signal processing calibration purposes. Calibrants are preferably utilized before a real sample is assessed. Calibrants can interface with a detection device using the same or a separate interface as the samples. In a preferred embodiment, calibrants are sprayed into a detection device using a second interface (e.g., second spray tip).

Polypeptide Detection

[0062] Detection devices can comprise of any device that is able to detect polypeptide presence and/or level, including for example, NMR, 2-D PAGE technology, Western blot technology, immunoanalysis technology and mass spectrometry. In a preferred embodiment, the system herein relies on a mass spectrometry to detect polypeptides present in a given sample. There are various forms of mass spectrometers that may be utilized.

[0063] In a preferred embodiment, an ESI-MS detection device is utilized. An ESI-MS combines the novelty of ESI with mass spectrometry. Furthermore, an ESI-MS preferably

utilizes a time-of-flight (TOF) mass spectrometry system. In TOF-MS, ions are generated by whatever ionization method is being employed and a voltage potential is applied. The potential extracts the ions from their source and accelerates them towards a detector. By measuring the time it takes the ions to travel a fixed distance, the mass of the ions can be calculated. TOF-MS can be set up to have an orthogonal-acceleration (OA). OA-TOF-MS are advantageous and preferred over conventional on-axis TOF because they have better spectral resolution and duty cycle. OA-TOF-MS also has the ability to obtain spectra at a relatively high speed. In addition to the MS systems disclosed above, other forms of ESI-MS include quadrupole mass spectrometry, ion trap mass spectrometry, and Fourier transform ion cyclotron resonance (FTICR-MS).

[0064] Quadrupole mass spectrometry consists of four parallel metal rods arranged in four quadrants (one rod in each quadrant). Two opposite rods have a positive applied potential and the other two rods have a negative potential. The applied voltages affect the trajectory of the ions traveling down the flight path. Only ions of a certain mass-to-charge ratio pass through the quadrupole filter and all other ions are thrown out of their original path. A mass spectrum is obtained by monitoring the ions passing through the quadrupole filter as the voltages on the rods are varied.

[0065] Ion trap mass spectrometry uses three electrodes to trap ions in a small volume. The mass analyzer consists of a ring electrode separating two hemispherical electrodes. A mass spectrum is obtained by changing the electrode voltages to eject the ions from the trap. The advantages of the ion-trap mass spectrometer include compact size, and the ability to trap and accumulate ions to increase the signal-to-noise ratio of a measurement

[0066] FTICR mass spectrometry is a mass spectrometric technique that is based upon an ion's motion in a magnetic field. Once an ion is formed, it eventually finds itself in the cell of the instrument, which is situated in a homogenous region of a large magnet. The ions are constrained in the XY plane by the magnetic field and undergo a circular orbit. The mass of the ion can now be determined based on the cyclotron frequency of the ion in the cell.

[0067] In a preferred embodiment, the system herein employs a TOF mass spectrometer, or more preferably, an ESI-TOF-MS, or more preferably an OA-TOF-MS, or more preferably a mass spectrometer having a dual ion funnel to support dynamic switching between multiple quadrupoles in series, the second of which can be used to dynamically filter ions by mass in

real time. In preferred embodiments, the detection device yields a spectrum of at least 1, more preferably 10, more preferably 20, or more preferably 50 spectra per second.

[0068] The detection device preferably interfaces with a separation/preparation device or microfluidic device, which allows for quick assaying of many of the polypeptides in a sample, or more preferably, most or all of the polypeptides in a sample. Preferably, a mass spectrometer is utilized that will accept a continuous sample stream for analysis and provide high sensitivity throughout the detection process (e.g., an ESI-MS). In another preferred embodiment, a mass spectrometer interfaces with one or more electrosprays, two or more electrosprays, three or more electrosprays or four or more electrosprays. Such electrosprays can originate from a single or multiple microfluidic devices.

[0069] The detection system utilized preferably allows for the capture and measurement of most or all of the polypeptides that are introduced into the detection device. It is preferable that one can observe polypeptides with high information-content that are only present at low concentrations. By contrast, it is preferable to remove those in advance that are, for example, common to all cells, especially those in high abundance.

Signal Processing/Pattern Recognition

[0070] The output from a detection device can then be processed, stored, and further analyzed or assayed using a bio-informatics system. A bio-informatics system can include one or more of the following: a computer; a plurality of computers connected to a network; a signal processing tool(s); a pattern recognition tool(s); and optionally a tool(s) to control flow rate for sample preparation, separation, and detection.

[0071] Data processing utilizes mathematical foundations. Generally, dynamic programming is preferably used to align a separation axis with a standard separation profile. Furthermore, intensities may be normalized, preferably by dividing by the total ion current of a spectrum. The data sets are then fitted using wavelets or other methods that are specifically designed for separation and mass spectrometer data. Data processing preferably filters out some of the noise and reduces spectrum dimensionality. This allows the system to identify the more highly predictive patterns.

[0072] In some embodiments, data processing may also involve the calibration of a mass-axis using linear correction determined by the calibrants. Calibration can take place prior to any sample detection; after sample detection; or in recurring intervals, for example.

[0073] Following data processing, pattern recognition tools are utilized to identify subtle differences between phenotypic states. Pattern recognition tools are based on a combination of statistical and computer scientific approaches, which provide dimensionality reduction. Such tools are scalable.

[0074] It is to be understood that the above embodiments are illustrative and not restrictive. The scope of the invention should be determined with respect to the scope of the appended claims, along with their full scope of equivalents.